# Genotypic and phenotypic consequences of domestication in dogs

**Authors:** Sweetalana[1], Jazlyn A Mooney[2,*] and Zachary A Szpiech[1,3,*]

[1]Department of Biology, Pennsylvania State University, University Park, PA, USA

[2]Department of Quantitative and Computational Biology, University of Southern California, Los Angeles, CA, USA

[3]Institute for Computational and Data Sciences, Pennsylvania State University, USA

Email for all authors: sweetalana2000@gmail.com, jazlyn.mooney@gmail.com, szpiech@psu.edu

*Indicates co-correspondence

Mailing Addresses with zip codes for all authors:

[1]514A Wartik Laboratory, University Park, PA, USA 16802

[2]1050 Childs Way, Ray R. Irani Hall, Los Angeles, CA, USA, 90089

ORCID:
Sweetalana: 0009-0000-5450-5934
Zachary A Szpiech: 0000-0001-6372-8224
Jazlyn A Mooney: 0000-0002-2369-0855

Short running title: Domestication shapes dog genetics and traits

Keywords: runs of homozygosity, dogs, inbreeding, GWAS, complex traits

## Abstract

Runs of homozygosity (ROH) are genomic regions that arise when two copies of an identical ancestral haplotype are inherited from parents with a recent common ancestor. In this study, we performed a novel comprehensive analysis to infer genetic diversity among dogs and quantified the association between ROH and non-disease phenotypes. We found distinct patterns of genetic diversity across clades of breed dogs and elevated levels of long ROH, compared to non-domesticated dogs. These high levels of $F_{ROH}$ (inbreeding coefficient) are a consequence of recent inbreeding among domesticated dogs during breed establishment. We identified statistically significant associations between $F_{ROH}$ and height, weight, lifespan, muscled, white head, white chest, furnish, and length of fur. After correcting for population structure, we identified more than 45 genes across the three examined quantitative traits that exceeded the threshold for suggestive significance, indicating significant polygenic inheritance for the complex quantitative phenotypes in dogs.


## Introduction

Dogs are one of the most common domesticated animals and evolved ~15,000 years ago from wolves (Vilà et al. 1997; Horard-Herbin, Tresset, and Vigne 2014; A. H. Freedman et al. 2014). Over time, humans performed tightly controlled selective breeding that resulted in various phenotypic traits (Vaysse et al. 2011). This intense selection led to the emergence of hundreds of new breeds and dramatic loss of genetic variation (Mastrangelo et al. 2018; Cecchi et al. 2013; Yordy et al. 2019). Due to breed establishment with a small number of founders and strong selection for morphology characteristics, high levels of inbreeding are observed across most domestic dog breeds (Bannasch et al. 2021). These high levels of inbreeding result in long

46 stretches of identical-by-descent (IBD) haplotypes in individual genomes, also known as runs of
47 homozygosity (ROH). Additionally, the process by which breeds were established resulted in a
48 unique genomic architecture, making testing phenotypic associations with multiple breeds
49 plausible (Plassais et al. 2019; Boyko 2011; Sams and Boyko 2019).

50 In humans, ROH are associated with numerous complex disease and non-disease
51 phenotypes (Clark et al. 2019; Johnson, Evans, and Keller 2018; Moreno-Grau et al. 2021; Keller
52 et al. 2012; Bacolod et al. 2008), and in dogs ROH have been associated with disease phenotypes
53 (Mooney, Yohannes, and Lohmueller 2021; Sams and Boyko 2019; Boyko 2011). Previous
54 studies have shown the influence of genetic variants in different quantitative traits such as leg
55 length, wrinkled skin, coat color, hair length, and skeletal shape (Sutter et al. 2007; Boyko et al.
56 2010; Parker et al. 2009; Cadieu et al. 2009; Candille et al. 2007). However, associations with
57 non-disease complex phenotypes in dogs are relatively unexplored (Hayward et al. 2016). To
58 determine the relationship between ROH and non-disease phenotypes in dogs, we analyzed 556
59 canid whole-genome sequences and characterized the genomic distribution of ROH. After
60 accounting for breed structure, we test for associations between the total ROH coverage of the
61 genome (as measured by $F_{ROH}$) among 13 breed groups with 13 phenotypes. We observe a
62 significant association with height, weight, lifespan, muscled, White chest, White head and
63 furnish. Finally, we perform an ROH-mapping genome-wide association study and identify
64 multiple genetic variants associated with three quantitative traits, namely height, weight, and
65 lifespan. Understanding the distribution of ROH and its association with non-disease phenotypes
66 provides insights on complex phenotypic architectures. Further, our work could help breeders
67 and veterinarians to evaluate biological mechanisms underlying complex phenotype variation in
68 dogs.

# Methods

## Data filtering and categorizations

In this study, we utilized 722 canid whole genomes from (Plassais et al. 2019) (NCBI accession number: PRJNA448733) that were filtered using BCFtools (Danecek et al. 2021). We retained only biallelic SNPs and sites of genotype quality >= 20 along with missing genotype rate of less than 10%. We filtered to remove variants from sex chromosomes, which resulted in 4,053,761 biallelic SNPs for all 38 pairs of autosomal chromosomes. Next, we categorized all individuals based on their Breed/Common name (Plassais et al. 2019) using both the neighbor joining trees from (vonHoldt et al. 2010) and the American Kennel Club (AKC) classifications. Lastly, we removed all wild canids, unknown breed status individuals, mixed breed individuals, and individuals for which breed information could not be found. This resulted in 13 breed groups namely, Ancient Spitz (31 individuals and 11 breeds), Herding (76 individuals and 15 breeds), Mastiff-like (65 individuals and 15 breeds), Retrievers (50 individuals and 6 breeds), Scent Hound (23 individuals and 10 breeds), Small Terrier (102 individuals and 8 breeds), Terriers (16 individuals and 7 breeds), Spaniels (22 individuals and 8 breeds), Toy Dogs (18 individuals and 7 breeds), Working Dogs ( 44 individuals and 10 breeds), Sight Hound (21 individuals and 9 breeds), Village Dogs (75 individuals and 13 regions), and Chinese Indigenous dogs (15 individuals). Overall, our working dataset included 558 individuals and 106 different breeds (Table S1).

## Calling Runs of Homozygosity

To call runs of homozygosity (ROH) we used GARLIC v1.6.0a, a likelihood-based ROH inference method (Szpiech, Blant, and Pemberton 2017), which uses a logarithm of odds (LOD) score measure of autozygosity applied in a sliding window for the entire genome (Pemberton et al. 2012; Szpiech, Blant, and Pemberton 2017). GARLIC requires input data with genotype and sample information in the form of TPED and TFAM files, which were generated using PLINK 2.0 (C. C. Chang et al. 2015). Moreover, we generated a TGLS file to obtain per-genotype likelihoods. We employed genotype likelihood data in the form of GQ to account for phred-scaled probability being incorrect. Next, we allowed GARLIC to choose the best window size based on SNP density, which starts at the window size of 100 SNPs and increases by the step size of 10 SNPs. GARLIC provides in-built ROH length classification for which we chose 6 classes: <1 Mbp, 1 Mbp -2 Mbp, 2 Mbp - 3 Mbp, 3 Mbp - 4 Mbp, 4 Mbp - 5 Mbp and >5 Mbp. To account for sample size differences in allele frequency estimation, we set the number of resamples to 40. All other flags were set to default, and the following parameters were used: --auto-winsize --auto-overlap-frac --winsize 100 --centromere centromeres_file.txt --size-bounds 1000000 2000000 3000000 4000000 5000000 --tped input_file.tped --tfam input_file.tfam --tgls input_file.tgls –gl-type GQ --resample 40 --out – output_file. To eliminate segments that were noisy, very short, and common, we filtered to retain ROH segments longer than 0.5 Mbp.

We performed independent ROH calling on breeds with at least 10 individuals. Breeds with fewer than 10 individuals were combined in their respective breed groups (Table S2). Thereafter, the ROH calling of all breeds irrespective of the individual numbers were combined into their breed groups for further analysis. For each individual, we computed the associations

between total number of runs of homozygosity ($n_{ROH}$) and total length of runs of homozygosity ($s_{ROH}$). For each breed group, we computed the arithmetic mean and range of both total $n_{ROH}$ and total $s_{ROH}$. For each individual and breed group, we computed the arithmetic mean of both nROH and sROH for each ROH size classification separately. The longest total sROH ( 2141.22 Mbp) and the highest total nROH (1039) were identified in two individuals - PER00747 and PER00393, respectively - from the small terrier Yorkshire breed. PER00747 also had a depth of the sequencing coverage across the entire genome of approximately 2.03, and produced inflated values of sROH. Thus, we dropped PER00747 and PER00393 given coverage was likely not sufficient for yielding reliable results.

## Inbreeding coefficient

For each sample and breed group, we computed a genomic inbreeding coefficient, $F_{ROH}$. $F_{ROH}$ is defined as the fraction of autosomal genome in ROH regions (McQuillan et al. 2009):

$$F_{ROH} = \frac{s_{ROH}}{L_{Auto}}$$

Where $s_{ROH}$ is the sum of length of runs of homozygosity in an individual's genome and $L_{Auto}$ is the length of autosomal genome in base pairs. Autosomal genome length of 2,203,764,842 base pairs was used since variants were called with the CanFam3.1 (NCBI RefSeq assembly : GCF_000002285.3) reference genome.

## ROH Sharing Matrix

We generated an ROH sharing matrix, a square matrix that represents the sum of overlaps of runs of homozygosity between two individuals, to account for shared ROH patterns among breeds. For this, we first generated individual ROH bed files for each individual. BEDTools

133 (Quinlan and Hall 2010) was used to identify those overlapping regions between two sets of
134 ROH data. The following parameters were used: bedtools intersect -wao -a
135 Firstindividual_ROH_file.bed -b Secondindividual_ROH_file.bed.

## 136 Phenotype data

137 We updated the phenotypic data (Plassais et al. 2019) for 13 phenotypes (bulky, Drop ears,
138 furnish, hairless, height, large ears, length of fur, lifespan, Long legs, muscled, weight, White
139 chest, and White head). For height, weight and lifespan, we used the average values obtained
140 from the AKC ("American Kennel Club," n.d.) (Table S3). When the sex-specific information
141 was available, we incorporated that into the dataset. When the information was absent, we used
142 the same values for both males and females. We normalized all three continuous traits using min-
143 max scaling (Table S4) . The phenotypic dataset for height, weight, and lifespan was further
144 segmented into small and large categories based on the average weight (21.66kg) of all
145 individuals (Table S5, Table S6). Additionally, we averaged the values for the remaining 10
146 binary traits based on the provided phenotypic data within their respective breeds (Table S7)
147 Here, the binary trait observed in an individual was encoded as 1, whereas traits not observed in
148 an individual was encoded as 0. No phenotypic information was calculated for the Chinese
149 indigenous dogs and Village dogs.

## 150 Association tests

151 We computed the association between $F_{ROH}$ and 13 phenotypes. We used GMMAT, an R
152 package that performs association tests using GLMM, generalized linear mixed models (Chen et
153 al. 2016). To fit the GLMM, glmmkin was used and we examined continuous traits (height,

154 weight, and lifespan) as the quantitative phenotype traits, $F_{ROH}$ as the covariate and an ROH
155 sharing matrix as kins. We fit the model assuming a Gaussian distribution of the continuous
156 phenotypes and identity as the link function. For associations of height, weight, and lifespan with
157 $F_{ROH}$ we used the updated phenotypic values for all 466 individuals. We also used weight and
158 weight*$F_{ROH}$ as the covariates across all three segments of the data (All individuals, small
159 individuals, and large individuals) (Table S8). To fit the GLMM for the other 10 binary traits, we
160 used binomial distribution as the family and logit as the link function. Individuals with "NA"
161 binary phenotypic information were removed from analysis. Due to lack of phenotype
162 information, Chinese Indigenous dogs and Village dogs were excluded from the association tests.

163 **ROH-mapping GWAS**

164 To understand the biological mechanisms of how ROH-associated SNPs influence the
165 phenotypic traits, we performed Genome-Wide Association Studies (GWAS). We generated a
166 dataset that examined the occurrence of SNPs in ROH regions using BCFTools intersect and
167 BCFTools subtract (Danecek et al. 2021). Here, SNPs in the ROH regions were encoded as 1,
168 whereas SNPs outside ROH regions were encoded as 0. We removed sites that were not in the
169 ROH regions (coded as 0) for all individuals. Next, we fit separate GLMM models using
170 GMMAT for all three phenotypic traits (height, weight, and lifespan) where ROH associated
171 SNPs were used as the covariate, ROH sharing matrix as the kins, Gaussian distribution as the
172 family and identity as the link function (Chen et al. 2016). We also fit GLMM models for each
173 segment of the data (All individuals, small individuals, and large individuals) using weight and
174 weight*interaction as the covariates (Table S9). We used the option effectiveSize within the coda
175 R package (Plummer et al. 2006) to obtain the effective sample size adjusted for autocorrelation

between sites. To obtain the genome-wide significance and suggestive significance thresholds, we used the Bonferroni correction with respect to effective sample size (Li et al. 2011; Yuan et al. 2020; W. Yang et al. 2014):

$$Genome - Wide\ Significance\ Threshold = -log10(\frac{0.05}{Effective\ Sample\ Size})$$

$$Suggestive - Wide\ Significance\ Threshold = -log10(\frac{0.1}{Effective\ Sample\ Size})$$

To correct for population stratification, we calculated the genomic inflation factor, λ (van den Berg et al. 2019; Sofer et al. 2021). We corrected the log-scaled p-values by dividing them by the genomic inflation factor (Table S9). To visualize the GWAS summary statistics, we used an R package, qqman to create Q-Q and Manhattan plots (Turner 2018).

# Results

## ROH distribution

We used GARLIC to assign each individual's ROH into our pre-specified ROH length classes. This allowed us to explore the variance per-population within each class. Briefly, ROH were categorized into 6 length classes: Class A (0.5 Mbp - 1 Mbp), Class B ( 1 Mbp - 2 Mbp), Class C ( 2 Mbp - 3 Mbp), Class D ( 3 Mbp - 4 Mbp), Class E (4 Mbp - 5 Mbp), and Class F ( > 5Mbp). The mean $n_{ROH}$ and mean $s_{ROH}$ for 556 individuals was 37.06 and 94.17 Mbp, respectively. We also reported the mean $n_{ROH}$ and mean $s_{ROH}$ for each ROH length classification: Class A (99.44 and 68.39 Mbp), Class B (50.29 and 70.36 Mbp), Class C (21.60 and 52.81 Mbp), Class D (13.07 and 45.23 Mbp), Class E ( 9.21 and 41.15 Mbp), and Class F (28.77 and 287.05 Mbp) (Table 1 and Table 2). Class A (0.5 Mbp - 1 Mbp) are the smallest ROH, and Class F are the largest ROH ( > 5Mbp).

197 ***TABLE 1.*** *nROH across all breed groups and ROH length classification. The table represents*
198 *the range and mean ± standard deviation of the number of runs of homozygosity for each breed*
199 *group and ROH length classification.*

200

| Breed Group | Mean $n_{ROH}$ | Range $n_{ROH}$ | Mean $n_{ROH}$ for Class A (0.5-1 Mbp) | Mean $n_{ROH}$ for Class B (1 - 2 Mbp) | Mean $n_{ROH}$ for Class C (2 - 3 Mbp) | Mean $n_{ROH}$ for Class D (3 - 4 Mbp) | Mean $n_{ROH}$ for Class E (4 - 5 Mbp) | Mean $n_{ROH}$ for Class F (> 5 Mbp) |
|---|---|---|---|---|---|---|---|---|
| **Ancient Spitz** | 47.10 ± 58.20 | 3 - 338 | 133.42 ± 85.63 | 65.13 ± 40.00 | 24.71 ± 11.23 | 13.90 ± 4.31 | 9.87 ± 3.97 | 35.58 ± 18.29 |
| **Retrievers** | 38.83 ± 34.58 | 0 - 138 | 106.42 ± 13.75 | 53.96 ± 9.05 | 23.98 ± 6.39 | 13.1 ± 4.31 | 9.94 ± 3.32 | 26.3 ± 11.77 |
| **Mastiff Like** | 38.37 ± 26.65 | 0 - 117 | 78.6 ± 14.12 | 54 ± 17.53 | 27.46 ± 11.19 | 16.91 ± 7.55 | 12.18 ± 6.53 | 41.08 ± 19.83 |
| **Herding** | 37.88 ± 33.97 | 0 - 257 | 97.76 ± 32.29 | 48.47 ± 10.17 | 21 ± 7.22 | 13.91 ± 5.79 | 10.8 ± 5.51 | 35.32 ± 19.08 |
| **Scent Hound** | 40.25 ± 37.52 | 1 - 160 | 113.26 ± 21.76 | 53.65 ± 9.85 | 22.7 ± 6.35 | 13.3 ± 4.87 | 9.09 ± 4.92 | 27.48 ± 12.44 |
| **Small Terrier** | 40.93 ± 41.19 | 0 - 277 | 113.58 ± 39.05 | 60.27 ± 17.74 | 24.08 ± 9.18 | 14.31 ± 6.32 | 8.98 ± 4.61 | 24.38 ± 16 |
| **Terriers** | 30.56 ± 33.01 | 9 - 122 | 90.75 ± 15.75 | 56.25 ± 11.22 | 31.81 ± 7.75 | 18.88 ± 6.21 | 15.88 ± 4.66 | 53.75 ± 10.77 |
| **Spaniels** | 36.29 ± 35.23 | 5 - 144 | 101.18 ± 20.19 | 61 ± 9.23 | 28.73 ± 5.37 | 18.14 ± 5.35 | 14 ± 5.19 | 41.64 ± 14.62 |
| **Sight Hound** | 45.29 ± 48.55 | 0 - 192 | 138.67 ± 31.05 | 63.33 ± 20.63 | 21.67 ± 8.98 | 13.86 ± 7.57 | 8.19 ± 4.76 | 26 ± 15.86 |
| **Toy Dogs** | 41.19 ± 40.77 | 1 - 187 | 114.56 ± 31.06 | 56.17 ± 25.99 | 24.06 ± 13.48 | 14.78 ± 9.78 | 9.44 ± 6.46 | 28.11 ± 15.34 |
| **Working Dogs** | 38.29 ± 27.43 | 1 - 129 | 85.18 ± 16.07 | 50.02 ± 15.37 | 27.36 ± 12.12 | 17 ± 7.86 | 11.82 ± 5.51 | 38.34 ± 13.11 |
| **Village Dogs** | 20.5 ± 31.83 | 0 - 261 | 80.12 ± 33.69 | 23.79 ± 13.69 | 6.33 ± 5.77 | 3.51 ± 4.76 | 2.09 ± 2.64 | 7.15 ± 11.03 |
| **Chinese Indigenous Dogs** | 14.36 ± 22.91 | 0 - 103 | 61.53 ± 16.63 | 14.67 ± 5.79 | 4 ± 3.31 | 1.27 ± 1.29 | 1 ± 1.32 | 3.67 ± 5.37 |

201

202 ***TABLE 2.*** *sROH across all breed groups and ROH length classification. The table represents*
203 *the range and mean ± standard deviation of the sum of runs of homozygosity for each breed*
204 *group and ROH length classification.*

| Breed Group | Mean $s_{ROH}$ (Mbp) | Range $s_{ROH}$ (Mbp) | Mean $s_{ROH}$ for Class A (0.5-1 Mbp) | Mean $s_{ROH}$ for Class B (1-2 Mbp) | Mean $s_{ROH}$ for Class C (2-3 Mbp) | Mean $s_{ROH}$ for Class D (3-4 Mbp) | Mean $s_{ROH}$ for Class E (4-5 Mbp) | Mean $s_{ROH}$ for Class F (> 5 Mbp) |
|---|---|---|---|---|---|---|---|---|
| **Ancient Spitz** | 119.56 ± 162.09 | 13.73 - 1113.00 | 92.29 ± 59.41 | 90.67 ± 54.30 | 60.33 ± 27.20 | 48.30 ± | 44.14 ± 17.51 | 381.60 ± 255.7 |
| **Retrievers** | 89.28 ± 86.63 | 0 - 598.02 | 72.96 ± 9.44 | 75.50 ± 13.25 | 57.37 ± 16.00 | 45.22 ± 15.07 | 44.47 ± 14.90 | 240.17 ± 125.98 |
| **Mastiff Like** | 126.47 ± 176.91 | 0 - 1029.28 | 54.77 ± 9.71 | 76.97 ± 25.95 | 67.21 ± 27.55 | 58.57 ± 25.92 | 54.38 ± 29.38 | 446.89 ± 247.25 |
| **Herding** | 105.25 ± 139.34 | 0 - 861.74 | 67.24 ± 21.71 | 67.53 ± 14.78 | 51.27 ± 17.64 | 48.23 ± 20.12 | 48.26 ± 24.73 | 348.94 ± 206.94 |
| **Scent Hound** | 99.19 ± 113.14 | 4.59 - 946.54 | 77.78 ± 14.73 | 74.84 ± 15.22 | 55.26 ± 15.67 | 45.90 ± 16.82 | 40.62 ± 21.78 | 300.75 ± 159.58 |
| **Small Terrier** | 89.41 ± 96.13 | 0 - 690.33 | 78.20 ± 26.38 | 84.24 ± 24.77 | 59.07 ± 22.76 | 49.63 ± 22.10 | 40.18 ± 20.51 | 225.15 ± 170.89 |
| **Terriers** | 75.86 ± 116.57 | 28.77 - 951.29 | 63.93 ± 10.90 | 79.69 ± 16.00 | 77.98 ± 19.72 | 65.84 ± 21.84 | 71.06 ± 20.79 | 541.71 ± 161.42 |
| **Spaniels** | 94.03 ± 127.54 | 23.23 - 856.91 | 70.13 ± 12.83 | 85.21 ± 12.79 | 70.39 ± 13.56 | 62.52 ± 18.65 | 62.55 ± 23.16 | 427.24 ± 190.19 |
| **Sight Hound** | 94.65 ± 96.81 | 0 - 591.60 | 95.24 ± 21.50 | 86.87 ± 29.21 | 53.04 ± 22.07 | 47.98 ± 26.13 | 36.54 ± 21.32 | 248.26 ± 149.49 |
| **Toy Dogs** | 97.01 ± 107.57 | 4.12 - 602.43 | 78.92 ± 22.17 | 79.01 ± 36.61 | 58.31 ± 33.13 | 50.45 ± 33.72 | 42.33 ± 28.82 | 273.02 ± 162.01 |
| **Working Dogs** | 112.89 ± 130.13 | 4.74 - 658.42 | 58.67 ± 11.14 | 70.67 ± 22.08 | 66.54 ± 29.19 | 58.83 ± 27.51 | 52.76 ± 24.70 | 369.85 ± 139.08 |
| **Village Dogs** | 33.24 ± 59.75 | 0 - 554.84 | 54.07 ± 23.09 | 32.65 ± 19.59 | 15.30 ± 13.97 | 12.04 ± 16.34 | 9.35 ± 11.72 | 76.00 ± 127.67 |
| **Chinese Indigenous Dogs** | 19.82 ± 35.06 | 0 - 309.27 | 41.03 ± 11.34 | 20.00 ± 7.64 | 9.53 ± 7.68 | 4.27 ± 4.24 | 4.41 ± 5.87 | 39.67 ± 75.16 |

205

206 Domestic dog clades showed a distinct pattern of genetic diversity, with higher levels of
207 long ROH (>5 Mbp) when compared with Chinese Indigenous Dogs (light green) and Village
208 dogs (royal blue) (Figure 1 and Figure 2). The highest total $n_{ROH}$ (641) and the highest total $s_{ROH}$
209 (1643.98 Mbp) were both observed in individuals (Dingo01 and NGSD2, respectively) from the
210 Ancient Spitz breed group (Table S10). The maximum length of a single ROH (1113.00 Mbp)
211 was observed in a New Guinea Singing Dog (Table 2 and Table S10). Terriers had the smallest
212 mean $n_{ROH}$ and $s_{ROH}$ when compared with the remaining 11 breed groups (Table 1 and Table 2).

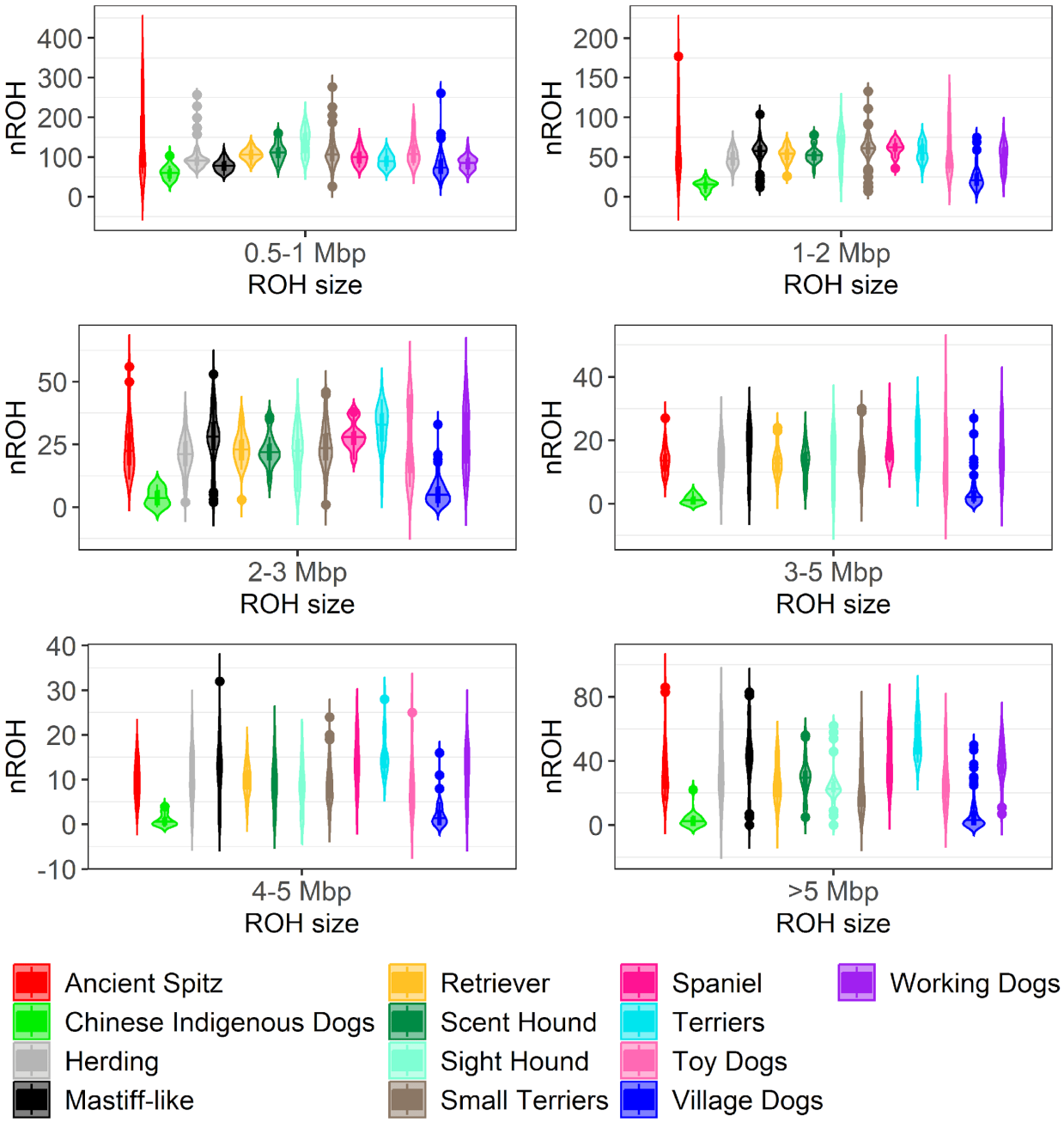



***FIGURE 1.*** *Relationship between number of runs of homozygosity (nROH) and ROH size class in Mbp for all breed groups. The x-axis represents the ROH size in Mbp. The y-axis represents the number of ROH. Each violin plot depicts the distribution of the total number of ROH within length ranges 0.5-1 Mbp, 1-2 Mbp, 2-3 Mbp, 3-4 Mbp, 4-5 Mbp and >5 Mbp for each breed group.*

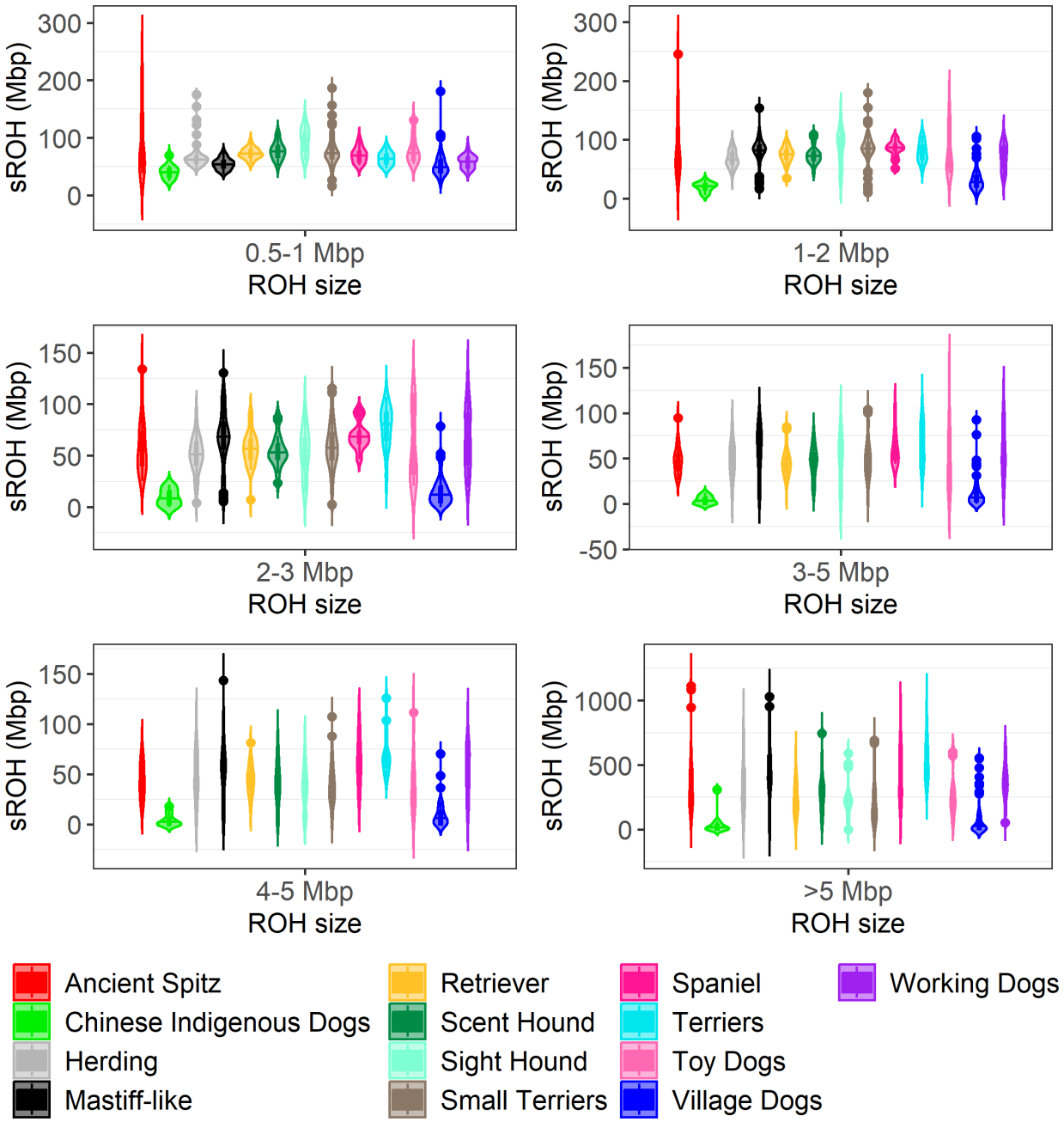



***FIGURE 2.*** *Relationship between sum of runs of homozygosity (sROH) and ROH size class in Mbp. The x-axis represents the ROH size in Mbp. The y-axis represents the sum of runs of homozygosity. Each violin plot depicts the distribution of the sum of ROH within length ranges 0.5-1 Mbp, 1-2 Mbp, 2-3 Mbp, 3-4 Mbp, 4-5 Mbp and >5 Mbp for each breed group.*

To further quantify the relationship between domesticated and non-domesticated populations, we examined $s_{ROH}$ and $n_{ROH}$ (Figure 3). Larger values $s_{ROH}$ have been shown in

227 populations that have experienced a bottleneck, while larger $n_{ROH}$ were found in small
228 populations (Ceballos et al. 2018). We observed larger values of mean $n_{ROH}$ and $s_{ROH}$ for the
229 domesticated breed dogs, relative to non-domesticated Chinese Indigenous Dogs and Village
230 Dogs (Figure 3). Further, the longest ROH were also observed in domestic dog clades. This
231 result corresponds with the higher degree of relatedness among all domestic breed group
232 individuals, and was expected given the known high levels of inbreeding during breed
233 establishment. Our results highlight Village and Chinese Indigenous dogs as more outbred
234 populations given lower ROH proportion when compared to breed dogs (Figure 3). We also
235 observed high levels of $n_{ROH}$ and $s_{ROH}$ within the Ancient Spitz clades. Our results suggest recent
236 bottlenecks in the Ancient Spitz clade's history.

237

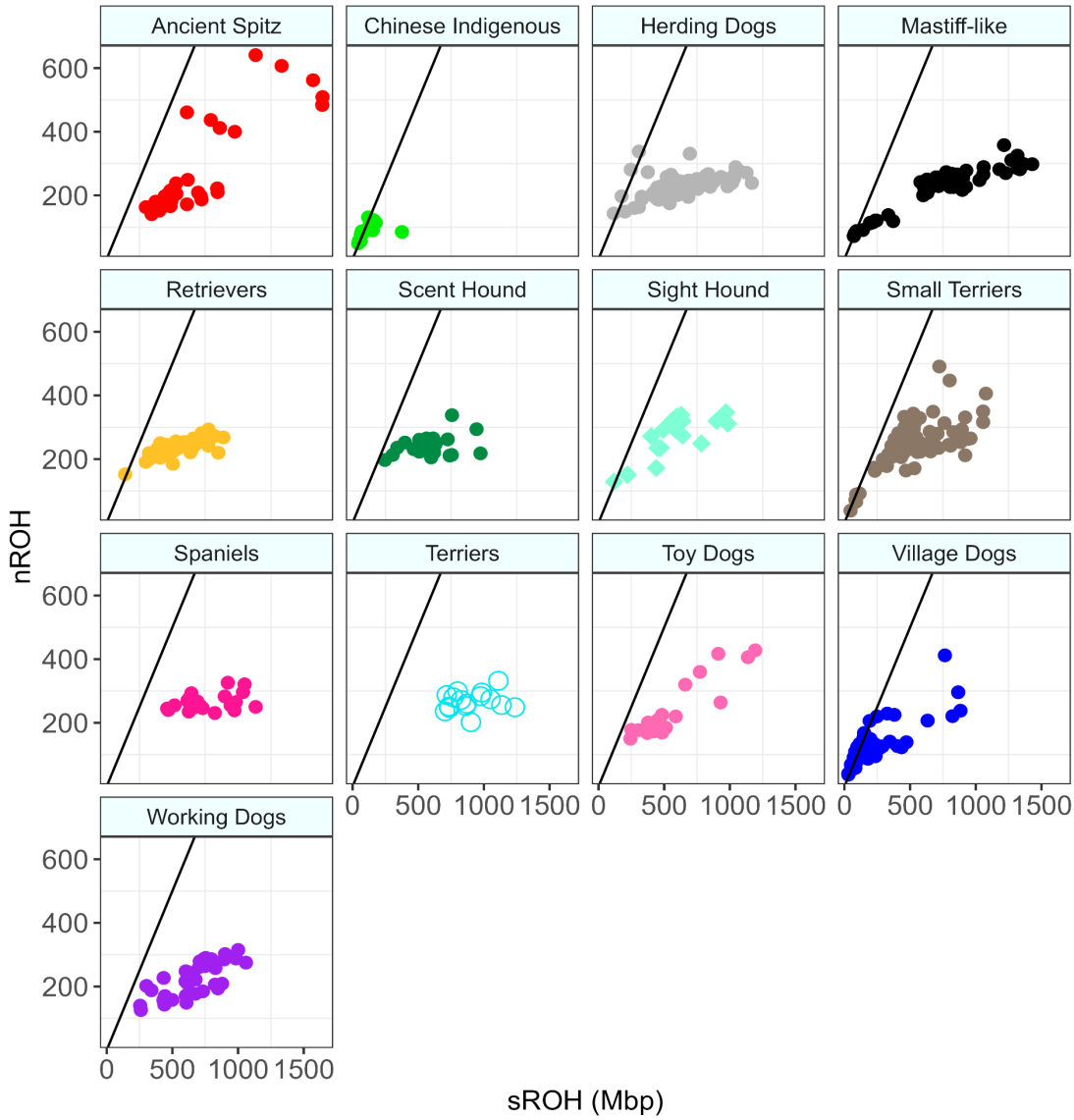



***FIGURE 3.*** *Relationship between sum of runs of homozygosity (sROH) and number of runs of homozygosity (nROH) for each breed group. The x-axis and the y-axis represents the sROH and nROH, respectively. The black line is y=x.*

Next, we examined the level of inbreeding within each clade by computing an inbreeding coefficient, $F_{ROH}$ for each individual and for the breed group (Table 3 and Table S11). To summarize the distribution of $F_{ROH}$ across all breed groups, we computed mean $F_{ROH}$ (Table 3). High levels of mean $F_{ROH}$ were observed in clades composed of breed dogs. $F_{ROH}$ values ranged from a mean of approximately 0.41 in Terriers to a mean of approximately 0.24 in Retrievers.

247 The lowest values were observed in Chinese Indigenous Dogs (mean ~0.05) and Village Dogs
248 (mean ~0.09). This result provides additional support for breed dogs sharing a more recent
249 common ancestor and more background relatedness relative to either the Village or Chinese
250 Indigenous dogs. Overall, our results from evaluating differences across $n_{ROH}$ and $s_{ROH}$ were
251 expected given what is known about the domestication of dogs, current populations of village
252 dogs, and the recent history of breed dogs and (Boyko 2011).

253 ***TABLE 3.*** *$F_{ROH}$ across all breed groups. The table represents the mean ± standard deviation of*
254 *$F_{ROH}$ for each breed group*

| Breed Group | Mean $F_{ROH}$ |
|---|---|
| **Terriers** | 0.41 ± 0.07 |
| **Spaniels** | 0.35 ± 0.09 |
| **Mastiff-Like** | 0.34 ± 0.15 |
| **Ancient Spitz** | 0.33 ± 0.17 |
| **Working Dog** | 0.31 ± 0.09 |
| **Herding** | 0.29 ± 0.11 |
| **Scent Hound** | 0.27 ± 0.08 |
| **Toy Dogs** | 0.26 ± 0.13 |
| **Sight Hound** | 0.26 ± 0.10 |
| **Small Terrier** | 0.24 ± 0.10 |
| **Retriever** | 0.24 ± 0.07 |
| **Village Dogs** | 0.09 ± 0.08 |
| **Chinese Indigenous Dogs** | 0.05 ± 0.04 |

255

256

## Association between $F_{ROH}$ and phenotypic traits

There exists a gap in the literature with respect to exploring the relationship between ROH and quantitative non-disease phenotypes. Thus, we next explored the relationship between $F_{ROH}$ and all 13 phenotypes (bulky, Drop ears, furnish, hairless, height, large ears, length of fur, lifespan, long legs, muscled, weight, White chest, and White head). Given that the effect of breed structure is particularly strong in dogs we and we were associating a trait with FROH we utilized with a kinship matrix quantified by pairwise ROH sharing between individuals, following the approach in Mooney et al. 2021. We identified three phenotypes for all individuals with significant associations: height ($\beta = 0.31$ and $p = 4.05 \times 10^{-3}$), weight ($\beta = 0.19$, $p = 1.19 \times 10^{-2}$), and lifespan ($\beta = -0.14$ and $p = 4.44 \times 10^{-2}$) (Figure 4). Across breed dogs, for every 1% increase in $F_{ROH}$, height increases by 0.192 cm, weight increases by 0.14 kg and lifespan decreases by 0.014 years. Statistical analysis including normalized coefficient Beta (β), along with its associated p-values and confidence intervals for different variations of models for quantitative traits (height, weight, lifespan) are shown in Table S8.

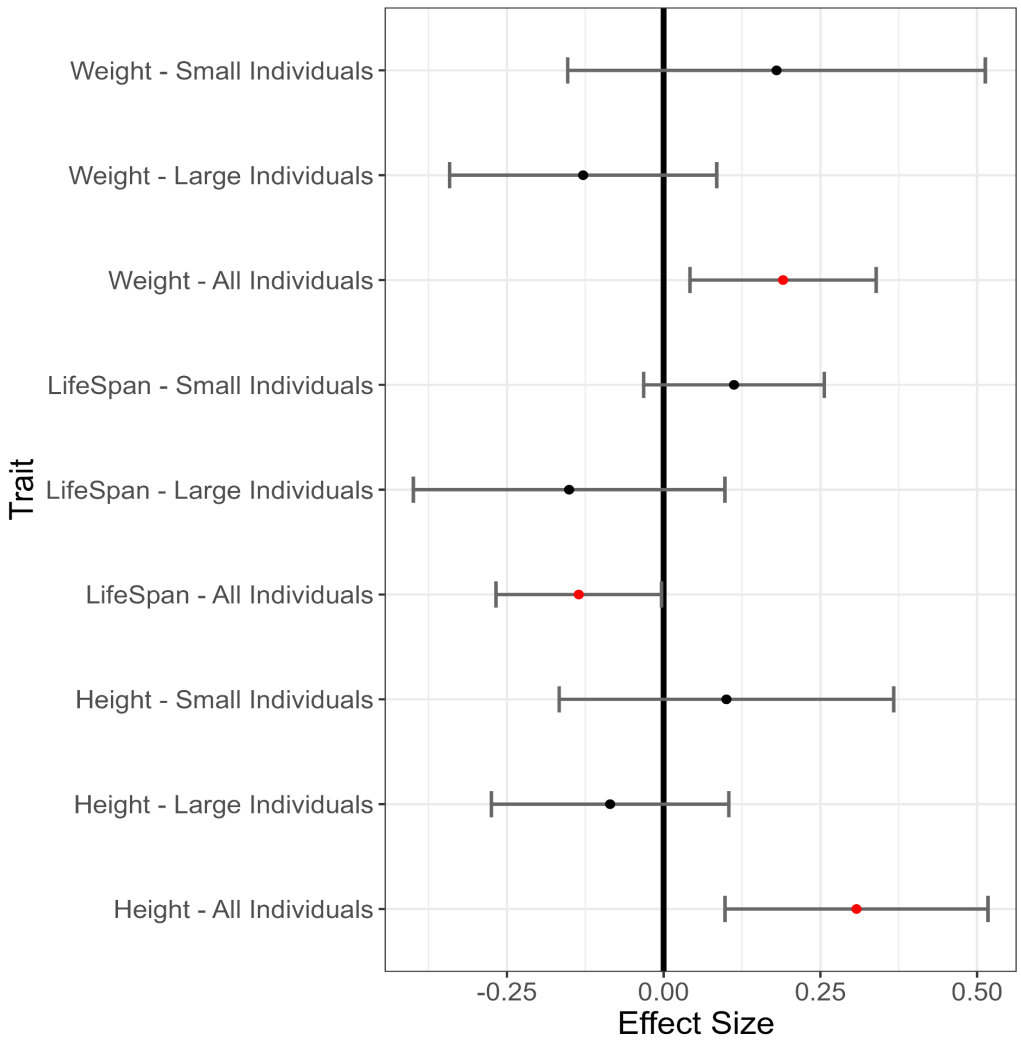



***FIGURE 4.*** *Relationship between effect size and quantitative phenotypic traits. The y-axis represents the phenotypes for which the associations were tested. The x-axis represents the normalized Beta coefficients (β) for their respective traits. A significant effect of FROH on a phenotype (nominal $p < 0.05$) is indicated with a red point. An effect size of greater than 0 indicates an increase of that trait with respect to $F_{ROH}$, and less than 0 represents the converse.*

Next, we determined whether there was a relationship between $F_{ROH}$ and the binary phenotypes. We observed significant associations between $F_{ROH}$ and 5 of the binary phenotypes: muscled ($\beta = 7.96$ and $p = 9.68 \times 10^{-11}$) , White chest ($\beta = 3.49$ and $p = 5.36 \times 10^{-5}$), White head ($\beta = 4.26$ and $p = 1.55 \times 10^{-4}$), length of fur ($\beta = -1.88$ and $p = 1.27 \times 10^{-2}$), and furnish

($\beta$ = -2.50 and p = $8.24 \times 10^{-3}$) (Figure 5 and Table S12). Across all breed groups, as $F_{ROH}$ increases, White Chest, White Head and muscled phenotypes also increase, whereas length of fur and furnish phenotypes decreases.

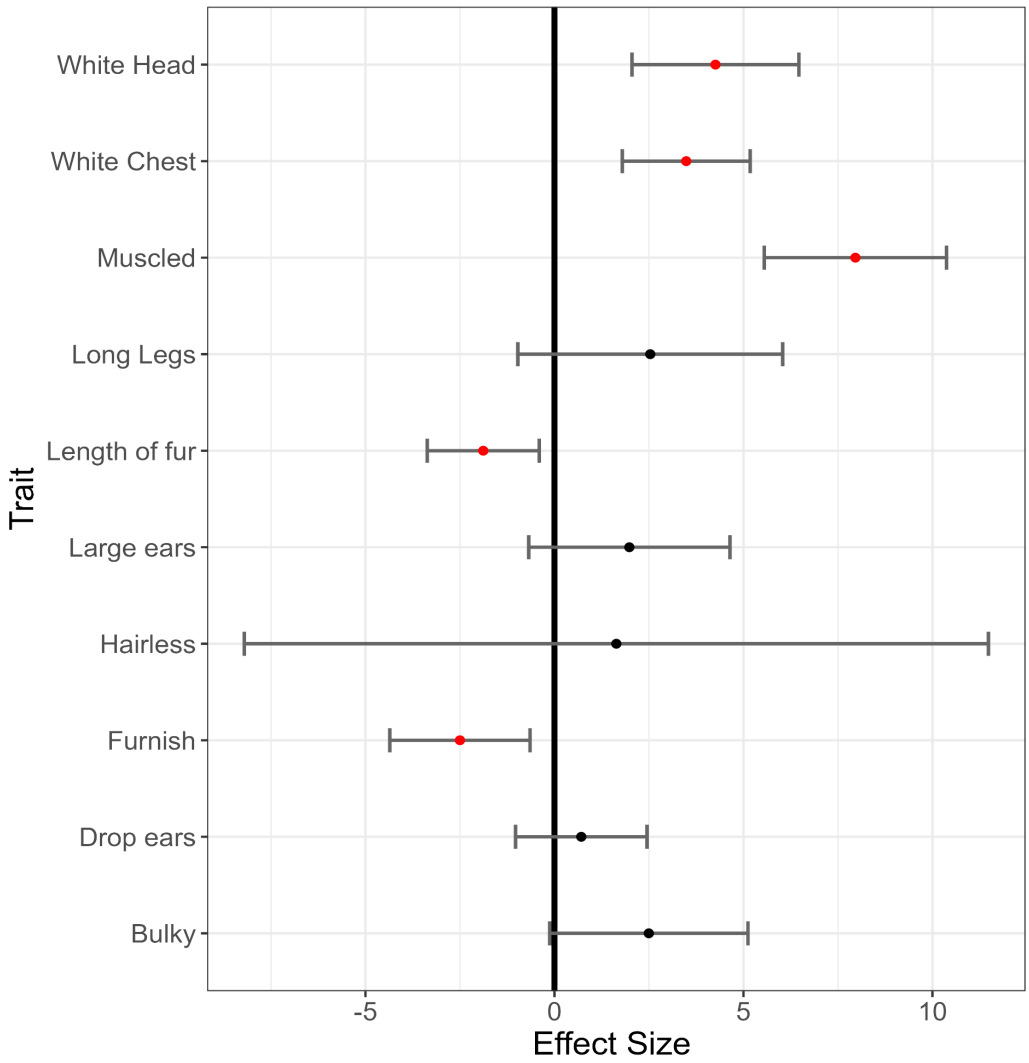



***FIGURE 5.*** *Relationship between effect size and categorical phenotypic traits. The y-axis represents the phenotypes for which the associations were tested. The x-axis represents the Beta coefficients (β) for their respective traits. A significant effect of $F_{ROH}$ on a phenotype (nominal p < 0.05) is indicated with a red point. An effect size of greater than 0 indicates an increase of that trait with respect to $F_{ROH}$, and less than 0 represents the converse.*

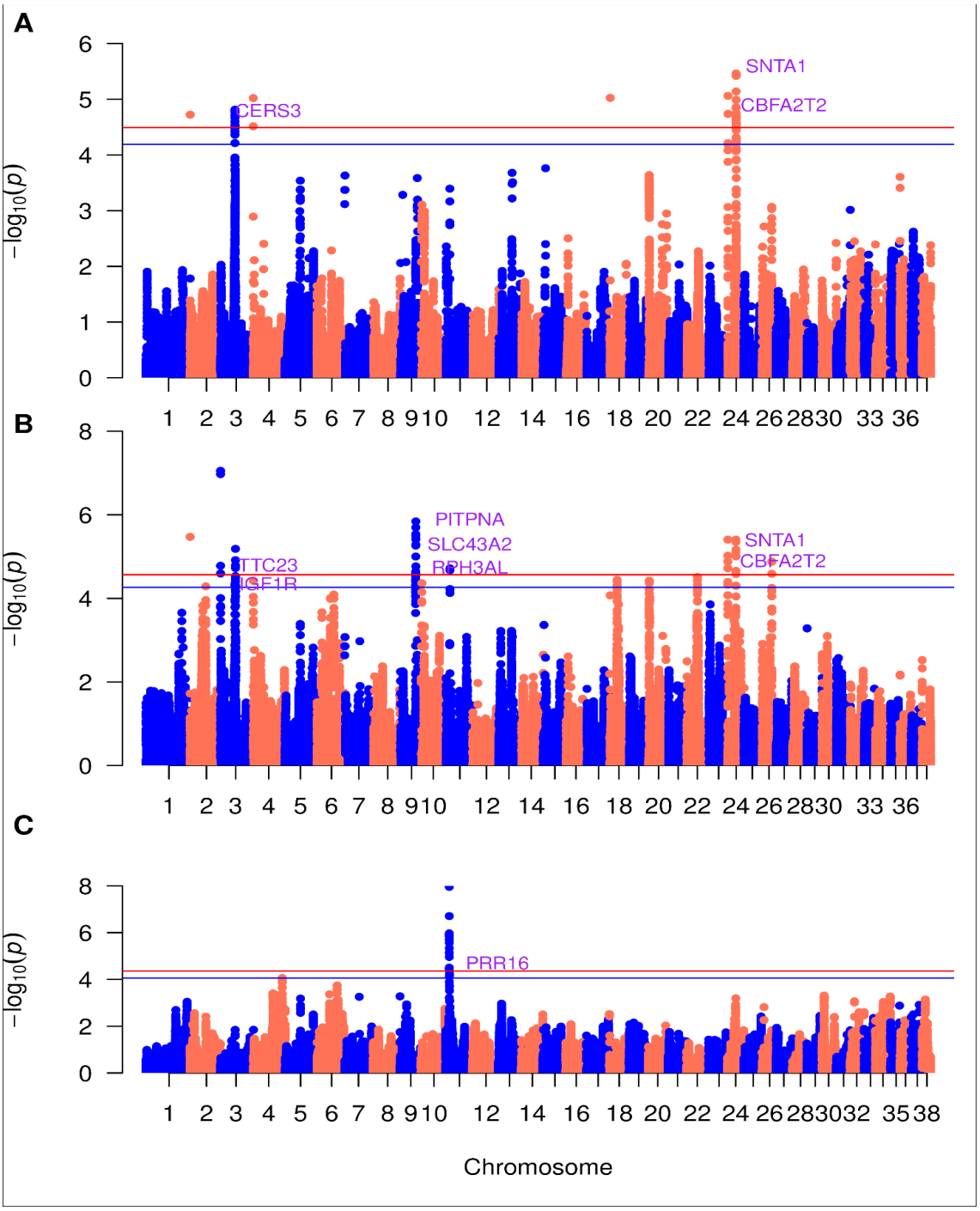
293 **ROH-mapping GWAS**

294 To identify ROH-associated genomic regions that influence life span, weight, or height, a
295 GLMM-based GWAS was performed (Figure 6, Figure 7, Figure 8, and Figure S1-S9). Out of
296 the 157,832 SNPs for height, 27 SNPs were observed to be above SWS (Suggestive Wide
297 Significance) threshold and 18 SNPs were observed to be above GWS (Genome Wide
298 Significance) threshold when tested for all individuals. A single SNP, chr24:22647289, was
299 detected to have the strongest associated genomic signal with height (p = $3.46 \times 10^{-6}$)
300 corresponding to SNTA1 (Figure 6A). Three genes, CERS3 (7 SNPs) on chromosome 3;
301 CBFA2T2 (16 SNPs) on chromosome 24; and SNTA1 (4 SNPs) on chromosome 24, were
302 observed above the SWS threshold, and contained multiple SNPs within their transcription
303 regions (Table S13). CERS3 was previously associated with Body Mass Index (BMI) in human
304 studies (Costa-Urrutia et al. 2019). Moreover, CERS3, SNTA1, and CBFA2T2 are known to
305 harbor common genetic variants that are found to be associated with height in humans (Kichaev
306 et al. 2019; Yengo et al. 2022).

307 For small individuals, 34 SNPs were observed to be above SWS threshold and 17 SNPs
308 were observed to be above GWS threshold when tested for association with height. A single
309 SNP, chr9:45711401 was detected to have the strongest associated genomic signal (p
310 $=1.43 \times 10^{-6}$) corresponding to the gene PITPNA (Figure 6B). On chromosome 3, CERS3 (3
311 SNPs), IGF1R (1 SNP), LRRC28 (1 SNP), and TTC23 (3 SNPs) contained common genetic
312 variants observed above SWS threshold. On chromosome 9, PITPNA (4 SNPs), SLC43A2 (2
313 SNPs), VPS53 (2 SNPs), and RPH3AL (4 SNPs) contained common genetic variants observed
314 above the SWS threshold. On chromosome 18, HGF (2 SNPs) was the only gene that contained

315 common genetic variants above the SWS threshold. On chromosome 22, COMMD6 (1 SNP)
316 was the only gene that contained a common genetic variant observed above the SWS threshold.
317 On chromosome 24, CBFA2T2 (7 SNPs) and SNTA1 (4 SNPs) contained common genetic
318 variants observed above the SWS threshold. Many of the genes (PITPNA, CERS3, LRRC28,
319 TTC23, HGF, CBFA2T2, SNTA1, SLC43A2, and VPS53) that we identified as reaching SWS
320 were previously identified as genes associated with height in humans (Yengo et al. 2022).
321 RPH3AL has been found to be associated with IGF1R in a network analysis with human patients
322 (Virakul et al. 2020). IGF1R has been previously associated with height in small dog breeds
323 (Hoopes et al. 2012). To our knowledge, no previous studies have identified an association
324 between height and COMMD6.

325 For large individuals, a single SNP, chr11:9943557 encoding for PRR16 was observed to
326 be above SWS threshold when tested for association with height (Figure 6C). PRR16 has also
327 been previously associated with height in humans (Yengo et al. 2022). An extended table
328 containing all the genes that were observed to be associated with height when using weight and
329 weight*interaction as the covariates are indicated in Table S13.



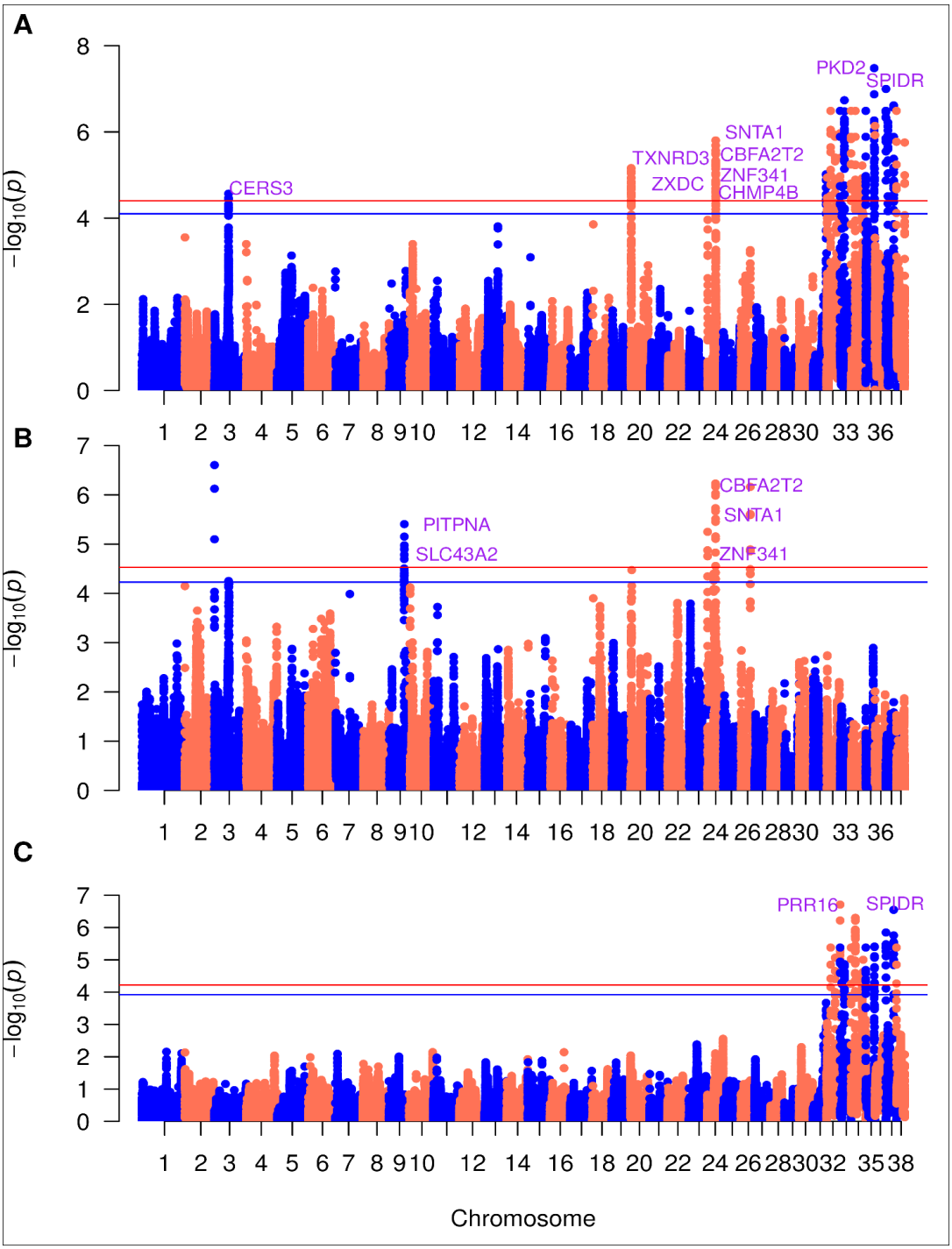

330

331 ***FIGURE 6.*** *Manhattan plots for GLMM-based GWAS for body height. A. All individuals B.*
332 *Small Individuals C. Large Individuals. The x-axis represents the genomic position. The y-axis*
333 *represents the log10 base transformed p-values. Single nucleotide polymorphisms are*
334 *represented by a single point. Red line indicates the suggestive wide significance (SWS)*
335 *threshold while the blue line indicates genome wide significance (GWS) threshold. Only genes*
336 *above the GWS threshold are indicated with the exception of IGF1R. All marked genes are not*
337 *drawn to scale.*

For weight, 66 SNPs were observed above SWS threshold with 45 SNPs above GWS threshold. The majority of SNPs (30) that were above the SWS threshold and found within genes were observed on chromosome 24 (Table S14). A single SNP, chr32:11413601, (p =$3.24 \times 10^{-7}$) with the strongest association with weight was observed within a gene that encodes for Protein Kinase D2, PKD2 (Figure 7A). Ablation of PKD2 was shown to be associated with reduced absorption of fats in the intestines of mice (Trujillo-Viera et al. 2021). PKD2 has also been linked to increase the capacity of an organism to absorb fat from a high fat diet, in other words, this gene potentially promotes obesity (Trujillo-Viera et al. 2021).

Other GWS genes when tested for all individuals with weight were SNTA1, CBFA2T2, TXNRD3, SPIDR, ZXDC, ZNF341, CERS3, and CHMP4B (Table S14). SNTA1 was found to be one of the closest associated genes with Live Body Weight (LBW) at 4 months in Karachai goats (Selionova et al. 2022). CBFA2T2 is a key regulator of adipogenesis which affects obesity when the gene expressions are altered (Matulewicz et al. 2017; Ali et al. 2013; Luo et al. 2020). TXNRD3, thioredoxin reductase 3 effects adipocyte differentiation through Wnt signaling pathway, a pathway that suppresses the adipogenic transcription factors (Martínez-Montes et al. 2016; Kipp et al. 2012; S. Kang et al. 2007). CERS3 was previously associated with Body Mass Index (BMI) in human studies (Costa-Urrutia et al. 2019). CHMP4B has been associated with human offspring birth weight (Warrington et al. 2019). Moreover, SPIDR was associated with increased total body fat and ZNF341 is associated with bone mineral content which is directly linked to body weight (Shapses and Sukumar 2012; Blake et al. 2021). ZNF341 has also been previously associated with weight in humans (Tachmazidou et al. 2017). However, to our knowledge, no previous studies have associated obesity/body mass with ZXDC and CHMP4B.

361 For small individuals, 31 SNPs were observed above SWS threshold and 19 SNPS above
362 GWS threshold. A single SNP, chr24:22660667, ($p = 5.88 \times 10^{-7}$) with the strongest
363 association with weight was observed within a gene that encodes for SNTA1 (Figure 7B).
364 ZNF341, PITPNA, CBFA2T2, and SLC43A2 were also observed above the GWS threshold.
365 Ablation of PITPNA is associated with decreased body weight in mice (Alb et al. 2003); and
366 ablation of SLC43A2 has been associated with 10% weight reduction in fetal mice (Alb et al.
367 2003; Guetg et al. 2015). In contrast, there were only two encoded genes in large individuals that
368 reached significance. PKD2 (chromosomes 32) and SPIDR (chromosome 35) were both
369 observed to be associated with weight above the SWS threshold (Figure 7C) .




370

371 ***FIGURE 7.*** *Manhattan plots for GLMM-based GWAS for body weight. A. All individuals B.*
372 *Small Individuals C. Large Individuals. The x-axis represents the genomic position. The y-axis*
373 *represents the log10 base transformed p-values. Single nucleotide polymorphisms are*
374 *represented by a single point. Red line indicates the suggestive wide significance (SWS)*
375 *threshold while the blue line indicates genome wide significance (GWS) threshold. All marked*
376 *genes are not drawn to scale.*

377 For lifespan, 10 SNPs were observed above the SWS threshold with only 5 SNPs above
378 the GWS threshold when tested across all individuals ( Figure 8A, Table S15). The strongest
379 association SNP, chr20:789204, (p = $2.21 \times 10^{-5}$) encoding for TXNRD3 was observed on
380 chromosome 20. TXNRD3 has been previously associated with lower survival rates for various
381 types of cancers, such as kidney renal cell carcinoma and pancreatic adenocarcinoma (Wu et al.
382 2021). However, to our knowledge, no previous studies have found that link lifespan with
383 TXNRD3. The next most significant SNP, chr33:18399003 (p = $2.68 \times 10^{-5}$) was observed in
384 the transcription region encoding for D3 dopamine receptor, DRD3. DRD3 is associated with
385 schizophrenia, resulting in loss of life expectancy (Y. Kang et al. 2023; Hjorthøj et al. 2017). We
386 observed another significant signal near, TIGIT, which has been associated with solid tumors
387 from lung, colon, breast, uterine and renal cancer, resulting in reduced longevity (Johnston et al.
388 2014; Chauvin et al. 2015). We also observed a signal above the SWS threshold on chromosome
389 32 in SEPT11, this gene was previously identified as a tumor-promoting gene with a predicted
390 poor prognosis (Fu et al. 2023). Likewise, ZXDC has been previously associated with enhanced
391 cervical cancer metastasis which has a significant negative impact on patient survival (Mao et al.
392 2023).

393 For small individuals, RUNX2 (4 SNPs), on chromosome 12, was the single gene that
394 was observed as a standout above the SWS threshold. (Figure 8B). RUNX2 has been previously
395 associated with multiple cancers such as breast cancer, prostate cancer, gastric cancer etc.,
396 resulting in poor prognosis (Zhao et al. 2021). For large individuals, TXNRD3, ZXDC,
397 CFAP100, and SLC41A3 were the observed genes above SWS threshold (Figure 8C). SLC41A3
398 has been previously associated with a probable indication for liver hepatocellular carcinoma
399 which is the second primary contributor of carcinoma-associated death (Zhao et al. 2021; Q.

400 Chang et al. 2021). To our knowledge, no previous studies have linked lifespan with CFAP100.
401 All associated genes that were observed with weight and weight*interaction covariates are listed
402 in Table S15.

403 Overall, our GWAS suggests that quantitative traits in dogs are polygenic with multiple
404 variants being associated with a trait. Our results also support previous research that has
405 demonstrated a much simpler genetic architecture of complex traits in breed dogs relative to
406 humans (Boyko et al. 2010). In the future, these genes could be used as candidates for further
407 functional exploration.

408

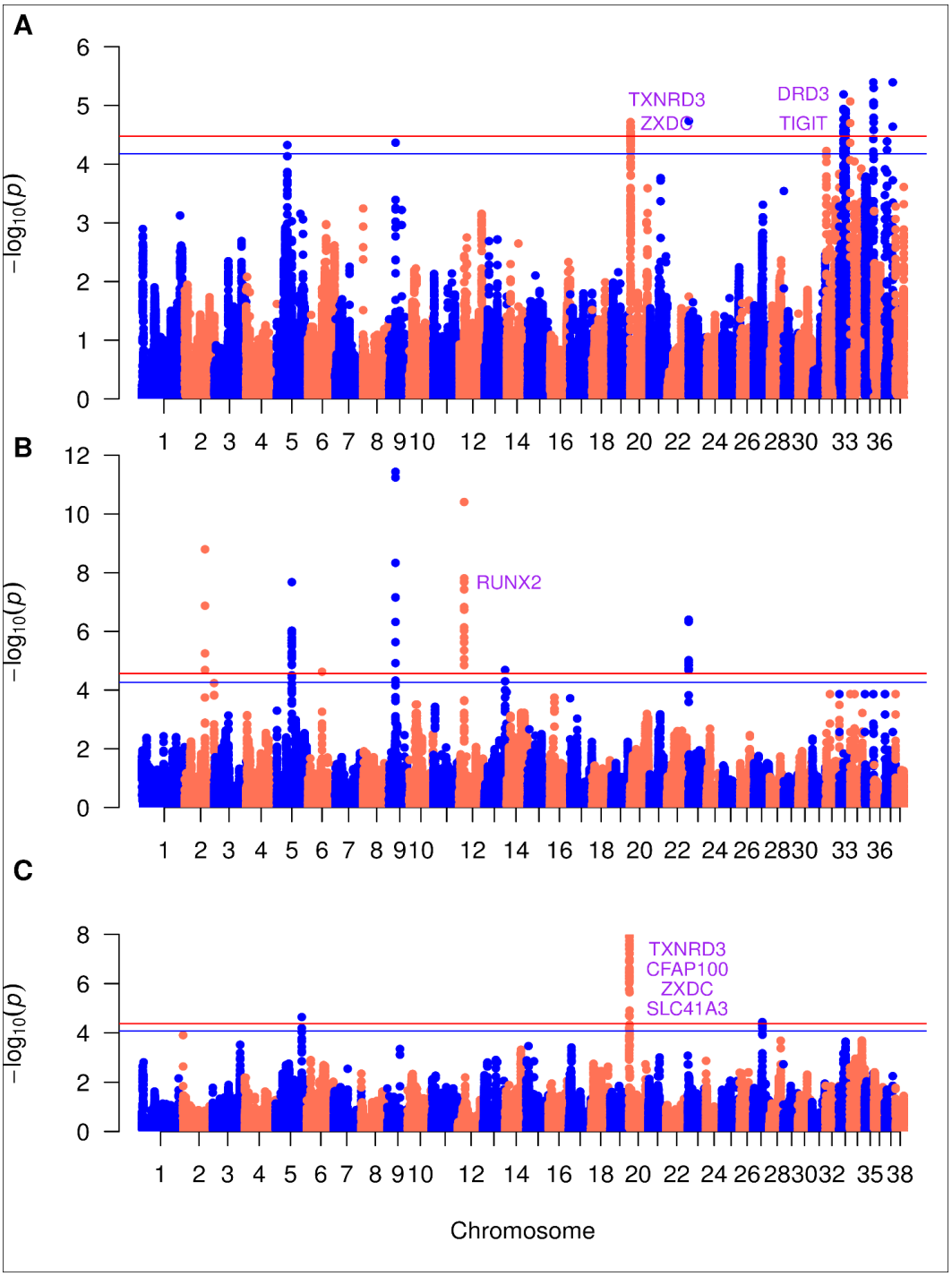



***FIGURE 8.*** *Manhattan plots for GLMM-based GWAS for lifespan. A. All individuals B. Small Individuals C. Large Individuals. The x-axis represents the genomic position. The y-axis represents the log10 base transformed p-values. Single nucleotide polymorphisms are represented by a single point. Red line indicates the suggestive wide significance (SWS) threshold while the blue line indicates genome wide significance (GWS) threshold. All marked genes are not drawn to scale.*

416 # Discussion

417 In sum, our work provides an in-depth analysis of ROH patterns in domesticated and non-
418 domesticated dogs. All breed dogs, which have experienced domestication, showed distinct
419 patterns of genetic diversity in comparison to Village dogs and Chinese Indigenous dogs. We
420 observe a larger fraction of the genome in long ROH (>5 Mbp) in breed dogs (Table 2), which
421 corresponds to small population sizes <10 generations ago (Purfield et al. 2017; Thompson
422 2013). This increase in the amount of the genome within a long ROH also corresponds with
423 higher values of $F_{ROH}$. Breed dogs also share most of their common genetic variation and lack
424 unique genetic diversity (Boyko et al. 2010). Lower genetic diversity and larger $F_{ROH}$ values
425 provide evidence for elevated levels of inbreeding and background relatedness. These results are
426 in line with findings from (Bannasch et al. 2021), and are indicative of the strong bottlenecks
427 experienced during breed formation.

428 Conversely, the non-domesticated Village dogs and Chinese Indigenous dogs have the
429 lowest values for mean ROH distribution. The Village dogs were sampled from 13 different
430 regions across the world and have been used as out groups in previous studies (Plassais et al.
431 2019). Their use as an out group is consistent with our results, where we observe lower values
432 for mean $n_{ROH}$ and mean $s_{ROH}$ (Table 1 and Table 2) in Village dogs relative to breed dogs.
433 Interestingly, we also observe lower levels of mean $n_{ROH}$ and mean $s_{ROH}$ in Chinese Indigenous
434 dogs. Our results are in line with the origins of the Chinese Indigenous dogs. Previous work
435 highlighted relaxed trait selection during establishment of Chinese Indigenous dog population
436 (Q. Yang et al. 2019), and we believe that this relaxed selection has resulted in overall higher
437 levels of genetic diversity.

438 We also examined the relationship between ROH and non-disease traits. We observed
439 that the relationship between height and $F_{ROH}$ is positive within all breed dogs (Figure 4). This
440 positive effect size (β) opposes recently published work in human populations (Swinford et al.
441 2023; Joshi et al. 2015; McQuillan et al. 2012). This may be a result of combining the β from
442 multiple sizes of dogs (i.e. large and small dogs). When we partition by weight, our results
443 correspond to previous studies on height in breed dogs (Hoopes et al. 2012). As one might
444 expect, we also find a positive β when testing for the relationship between $F_{ROH}$ and weight.
445 Lastly, we observed a negative β with respect to the relationship between $F_{ROH}$ and lifespan,
446 indicating an association with inbreeding and survival to old age. This result is consistent with
447 the findings in dogs where an association with disease phenotypes was observed (Mooney,
448 Yohannes, and Lohmueller 2021). Further, we also partitioned dogs by weight and observed a
449 trend similar to (Urfer et al. 2019), where larger dogs have a shorter lifespan and smaller dogs
450 have a longer lifespan, though our results were not significant (Table S8).

451 We identified multiple SNPs within genes that were associated with height, weight, and
452 lifespan (Table S13, Table S14, and Table S15). Many of the genes that were observed in our
453 study were not present in previous work, highlighting the necessity to enhance the correction of
454 existing genetic structure and explore complex trait architecture in domesticated species. In
455 humans, there have been 25,551 associations with height, 2,233 associations with weight, and
456 664 associations with lifespan as of 2023 (Sollis et al. 2023). After correcting for inflation, we
457 recover a total of 50 associations with height, 84 associations with weight, and 21 associations
458 with lifespan. This corresponds to previous work which has suggested dogs having a more
459 simplified complex trait architecture (Boyko et al. 2010; Adam H. Freedman, Lohmueller, and
460 Wayne 2016).

There some limitations in the work that we would like to highlight. First, the phenotypic values used in the analysis were calculated based on the average, restricting us to study associations within breeds. Secondly, breed group categorization based on both neighbor joining trees and AKC could lead to errors when grouping some of breeds. Additionally, we found that there was noise in association tests for the last few chromosomes. We believe the noise could be due to lower marker density on smaller chromosomes.

Despite these limitations, the significant relationship between $F_{ROH}$ and certain phenotypic traits suggest that inbreeding and artificial selection played a role in shaping complex non-disease traits within breed dogs. In sum, our work has further elucidated the genetic structure of quantitative trait architecture in dogs. We also highlight, how domestication and inbreeding, quantified in the form of ROH, affect the distribution of regions of the genome associated with complex non-disease phenotypes.

## Author Contributions

ZAS and JAM designed and supervised the research project. S performed the research and analyzed data. S along with the assistance of ZAS and JAM wrote the paper.

## Acknowledgements

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R35GM146926 (ZAS and S). This work was also supported by Eberly College of Science Startup Fund (ZAS and S). S would also like to thank the entire Szpiech Lab for their invaluable support. JAM was supported by the startup funds from

Dornsife College of Letters, Arts and Sciences through the Department of Quantitative and Computational Biology and the USC WiSE Gabilan Assistant Professorship.

## Data Accessibility Statement

Whole Genome Sequencing is available on NCBI (Accession number: PRJNA448733). The source phenotype data was obtained from (Plassais et al. 2019). All other data are contained within the article and its supplementary information. Code can be found on GitHub at (https://github.com/jaam92/QuantTraitInDogs.git).

# Bibliography


Alb, James G., Jr, Jorge D. Cortese, Scott E. Phillips, Roger L. Albin, Tim R. Nagy, Bruce A. Hamilton, and Vytas A. Bankaitis. 2003. “Mice Lacking Phosphatidylinositol Transfer Protein-Alpha Exhibit Spinocerebellar Degeneration, Intestinal and Hepatic Steatosis, and Hypoglycemia.” *The Journal of Biological Chemistry* 278 (35): 33501–18.

Ali, Aus Tariq, Warren E. Hochfeld, Renier Myburgh, and Michael S. Pepper. 2013. “Adipocyte and Adipogenesis.” *European Journal of Cell Biology* 92 (6-7): 229–36.

“American Kennel Club.” n.d. American Kennel Club. https://www.akc.org/.

Bacolod, Manny D., Gunter S. Schemmann, Shuang Wang, Richard Shattock, Sarah F. Giardina, Zhaoshi Zeng, Jinru Shia, et al. 2008. “The Signatures of Autozygosity among Patients with Colorectal Cancer.” *Cancer Research* 68 (8): 2610–21.

Bannasch, Danika, Thomas Famula, Jonas Donner, Heidi Anderson, Leena Honkanen, Kevin Batcher, Noa Safra, Sara Thomasy, and Robert Rebhun. 2021. “The Effect of Inbreeding, Body Size and Morphology on Health in Dog Breeds.” *Canine Medicine and Genetics* 8 (1): 12.

Berg, Sanne van den, Jérémie Vandenplas, Fred A. van Eeuwijk, Marcos S. Lopes, and Roel F. Veerkamp. 2019. “Significance Testing and Genomic Inflation Factor Using High-density Genotypes or Whole-genome Sequence Data.” *Journal of Animal Breeding and Genetics = Zeitschrift Fur Tierzuchtung Und Zuchtungsbiologie* 136 (6): 418.

Blake, Judith A., Richard Baldarelli, James A. Kadin, Joel E. Richardson, Cynthia L. Smith, and Carol J. Bult. 2021. “Mouse Genome Database (MGD): Knowledgebase for Mouse-Human Comparative Biology.” *Nucleic Acids Research* 49 (D1): D981–87.

Boyko, Adam R. 2011. “The Domestic Dog: Man’s Best Friend in the Genomic Era.” *Genome Biology* 12 (2): 216.

Boyko, Adam R., Pascale Quignon, Lin Li, Jeffrey J. Schoenebeck, Jeremiah D. Degenhardt, Kirk E. Lohmueller, Keyan Zhao, et al. 2010. “A Simple Genetic Architecture Underlies Morphological Variation in Dogs.” *PLoS Biology* 8 (8). https://doi.org/10.1371/journal.pbio.1000451.

Cadieu, E., M. W. Neff, P. Quignon, K. Walsh, K. Chase, H. G. Parker, B. M. Vonholdt, et al. 2009. “Coat Variation in the Domestic Dog Is Governed by Variants in Three Genes.” 10.1126/science.1177808.

Candille, S. I., C. B. Kaelin, B. M. Cattanach, B. Yu, D. A. Thompson, M. A. Nix, J. A. Kerns, S. M. Schmutz, G. L. Millhauser, and G. S. Barsh. 2007. “A -Defensin Mutation Causes Black Coat Color in Domestic Dogs.” 10.1126/science.1147880.

Ceballos, Francisco C., Peter K. Joshi, David W. Clark, Michèle Ramsay, and James F. Wilson. 2018. “Runs of Homozygosity: Windows into Population History and Trait Architecture.” *Nature Reviews. Genetics* 19 (4): 220–34.

Cecchi, Francesca, Gisella Paci, Andrea Spaterna, and Roberta Ciampolini. 2013. “Genetic Variability in Bracco Italiano Dog Breed Assessed by Pedigree Data.” *Italian Journal of Animal Science*, January. https://doi.org/10.4081/ijas.2013.e54.

Chalazan, Brandon, Denada Palm, Arvind Sridhar, Christina Lee, Maria Argos, Martha Daviglus, Jalees Rehman, Sreenivas Konda, and Dawood Darbar. 2021. “Common Genetic Variants Associated with Obesity in an African-American and Hispanic/Latino Population.” *PloS One* 16 (5). https://doi.org/10.1371/journal.pone.0250697.

Chang, Christopher C., Carson C. Chow, Laurent Cam Tellier, Shashaank Vattikuti, Shaun M. Purcell, and James J. Lee. 2015. “Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets.” *GigaScience* 4 (1): s13742–015 – 0047–0048.

Chang, Qimeng, Yayun Xu, Jianfa Wang, Hui Jing, Longhua Rao, Weiguo Tang, Ziping Zhang,

and Xubo Wu. 2021. “SLC41A3 Exhibits as a Carcinoma Biomarker and Promoter in Liver Hepatocellular Carcinoma.” *Computational and Mathematical Methods in Medicine* 2021. https://doi.org/10.1155/2021/8556888.

Chauvin, Joe-Marc, Ornella Pagliano, Julien Fourcade, Zhaojun Sun, Hong Wang, Cindy Sander, John M. Kirkwood, et al. 2015. “TIGIT and PD-1 Impair Tumor Antigen–specific CD8+ T Cells in Melanoma Patients.” *The Journal of Clinical Investigation* 125 (5): 2046–58.

Chen, Han, Chaolong Wang, Matthew P. Conomos, Adrienne M. Stilp, Zilin Li, Tamar Sofer, Adam A. Szpiro, et al. 2016. “Control for Population Structure and Relatedness for Binary Traits in Genetic Association Studies via Logistic Mixed Models.” *American Journal of Human Genetics* 98 (4): 653–66.

Clark, David W., Yukinori Okada, Kristjan H. S. Moore, Dan Mason, Nicola Pirastu, Ilaria Gandin, Hannele Mattsson, et al. 2019. “Associations of Autozygosity with a Broad Range of Human Phenotypes.” *Nature Communications* 10 (1): 4957.

Costa-Urrutia, Paula, Valentina Colistro, Angélica S. Jiménez-Osorio, Helios Cárdenas-Hernández, Jacqueline Solares-Tlapechco, Miryam Ramirez-Alcántara, Julio Iván Granados, and Martha E. Rodríguez-Arellano. 2019. “Genome-Wide Association Study of Body Mass Index and Body Fat in Mexican-Mestizo Children.” *Genes (Basel)* 10 (11): 945.

Danecek, Petr, James K. Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O. Pollard, Andrew Whitwham, et al. 2021. “Twelve Years of SAMtools and BCFtools.” *GigaScience* 10 (2): giab008.

Freedman, Adam H., Kirk E. Lohmueller, and Robert K. Wayne. 2016. “Evolutionary History, Selective Sweeps, and Deleterious Variation in the Dog.” *Annual Review of Ecology, Evolution, and Systematics* 47 (Volume 47, 2016): 73–96.

Freedman, A. H., I. Gronau, Schweizer RM, Vecchyo D. Ortega-Del, E. Han, P. M. Silva, M. Galaverni, et al. 2014. “Genome Sequencing Highlights the Dynamic Early History of Dogs.” *PLoS Genetics* 10 (1). https://doi.org/10.1371/journal.pgen.1004016.

Fu, Lisheng, Xiaoyan Wang, Ying Yang, Meihua Chen, Adilijiang Kuerban, Haojie Liu, Yiwei Dong, Qianqian Cai, Mingzhe Ma, and Xingzhong Wu. 2023. “Septin11 Promotes Hepatocellular Carcinoma Cell Motility by Activating RhoA to Regulate Cytoskeleton and Cell Adhesion.” *Cell Death & Disease* 14 (4): 280.

Galis, Frietson, Inke Van der Sluijs, Tom J. M. Van Dooren, Johan A. J. Metz, and Marc Nussbaumer. 2007. “Do Large Dogs Die Young?” *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution* 308 (2): 119–26.

Guetg, Adriano, Luca Mariotta, Lukas Bock, Brigitte Herzog, Ralph Fingerhut, Simone M. R. Camargo, and François Verrey. 2015. “Essential Amino Acid Transporter Lat4 (Slc43a2) Is Required for Mouse Development.” *The Journal of Physiology* 593 (5): 1273–89.

Hayward, Jessica J., Marta G. Castelhano, Kyle C. Oliveira, Elizabeth Corey, Cheryl Balkman, Tara L. Baxter, Margret L. Casal, et al. 2016. “Complex Disease and Phenotype Mapping in the Domestic Dog.” *Nature Communications* 7 (1): 1–11.

Hjorthøj, Carsten, Anne Emilie Stürup, John J. McGrath, and Merete Nordentoft. 2017. “Years of Potential Life Lost and Life Expectancy in Schizophrenia: A Systematic Review and Meta-Analysis.” *The Lancet Psychiatry* 4 (4): 295–301.

Hoopes, Barbara C., Maud Rimbault, David Liebers, Elaine A. Ostrander, and Nathan B. Sutter. 2012. “The Insulin-like Growth Factor 1 Receptor (IGF1R) Contributes to Reduced Size in Dogs.” *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 23 (11-12): 780–90.

Horard-Herbin, Marie-Pierre, Anne Tresset, and Jean-Denis Vigne. 2014. “Domestication and Uses of the Dog in Western Europe from the Paleolithic to the Iron Age.” *Animal Frontiers* 4 (3): 23–31.

Johnson, Emma C., Luke M. Evans, and Matthew C. Keller. 2018. “Relationships between

587 Estimated Autozygosity and Complex Traits in the UK Biobank." *PLoS Genetics* 14 (7):
588 e1007556.
589 Johnston, Robert J., Laetitia Comps-Agrar, Jason Hackney, Xin Yu, Mahrukh Huseni, Yagai
590 Yang, Summer Park, et al. 2014. "The Immunoreceptor TIGIT Regulates Antitumor and
591 Antiviral CD8+ T Cell Effector Function." *Cancer Cell* 26 (6): 923–37.
592 Joshi, P. K., T. Esko, H. Mattsson, N. Eklund, I. Gandin, T. Nutile, A. U. Jackson, et al. 2015.
593 "Directional Dominance on Stature and Cognition in Diverse Human Populations." *Nature*
594 523 (7561). https://doi.org/10.1038/nature14618.
595 Kang, Sona, Christina N. Bennett, Isabelle Gerin, Lauren A. Rapp, Kurt D. Hankenson, and
596 Ormond A. MacDougald. 2007. "Wnt Signaling Stimulates Osteoblastogenesis of
597 Mesenchymal Precursors by Suppressing CCAAT/Enhancer-Binding Protein α and
598 Peroxisome Proliferator-Activated Receptor γ." *The Journal of Biological Chemistry* 282
599 (19): 14515–24.
600 Kang, Yafei, Youming Zhang, Kexin Huang, and Zhenhong Wang. 2023. "The Genetic Influence
601 of the DRD3 rs6280 Polymorphism (Ser9Gly) on Functional Connectivity and Gray Matter
602 Volume of the Hippocampus in Patients with First-Episode, Drug-Naïve Schizophrenia."
603 *Behavioural Brain Research* 437 (February): 114124.
604 Keller, Matthew C., Matthew A. Simonson, Stephan Ripke, Ben M. Neale, Pablo V. Gejman,
605 Daniel P. Howrigan, Sang Hong Lee, et al. 2012. "Runs of Homozygosity Implicate
606 Autozygosity as a Schizophrenia Risk Factor." *PLoS Genetics* 8 (4): e1002656.
607 Kichaev, Gleb, Gaurav Bhatia, Po-Ru Loh, Steven Gazal, Kathryn Burch, Malika K. Freund,
608 Armin Schoech, Bogdan Pasaniuc, and Alkes L. Price. 2019. "Leveraging Polygenic
609 Functional Enrichment to Improve GWAS Power." *American Journal of Human Genetics*
610 104 (1): 65–75.
611 Kipp, Anna P., Mike F. Müller, Eva M. Göken, Stefanie Deubel, and Regina Brigelius-Flohé.
612 2012. "The Selenoproteins GPx2, TrxR2 and TrxR3 Are Regulated by Wnt Signalling in the
613 Intestinal Epithelium." *Biochimica et Biophysica Acta (BBA) - General Subjects* 1820 (10):
614 1588–96.
615 Li, Miao-Xin, Juilian M. Y. Yeung, Stacey S. Cherny, and Pak C. Sham. 2011. "Evaluating the
616 Effective Numbers of Independent Tests and Significant P-Value Thresholds in Commercial
617 Genotyping Arrays and Public Imputation Reference Datasets." *Human Genetics* 131 (5):
618 747–56.
619 Luo, Jun, Lei Dou, Zhengyan Yang, Zhi Zhou, and Hong Huang. 2020. "CBFA2T2 Promotes
620 Adipogenic Differentiation of Mesenchymal Stem Cells by Regulating CEBPA." *Biochemical*
621 *and Biophysical Research Communications* 529 (2): 133–39.
622 Mao, Yifang, Xingyu Jiang, Peng Guo, Ying Ouyang, Xiangfu Chen, Meng Xia, Lixin Wu, et al.
623 2023. "ZXDC Enhances Cervical Cancer Metastasis through IGF2BP3-Mediated Activation
624 of RhoA/ROCK Signaling." *iScience* 26 (8). https://doi.org/10.1016/j.isci.2023.107447.
625 Martínez-Montes, Angel M., Anixa Muiños-Bühl, Almudena Fernández, Josep M. Folch, Noelia
626 Ibáñez-Escriche, and Ana I. Fernández. 2016. "Deciphering the Regulation of Porcine
627 Genes Influencing Growth, Fatness and Yield-Related Traits through Genetical Genomics."
628 *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 28
629 (3): 130–42.
630 Mastrangelo, Salvatore, Filippo Biscarini, Barbara Auzino, Marco Ragatzu, Andrea Spaterna,
631 and Roberta Ciampolini. 2018. "Genome-Wide Diversity and Runs of Homozygosity in the
632 'Braque Français, Type Pyrénées' Dog Breed." *BMC Research Notes* 11 (1): 1–6.
633 Matulewicz, Natalia, Magdalena Stefanowicz, Agnieszka Nikolajuk, and Monika Karczewska-
634 Kupczewska. 2017. "Markers of Adipogenesis, but Not Inflammation, in Adipose Tissue Are
635 Independently Related to Insulin Sensitivity." *The Journal of Clinical Endocrinology and*
636 *Metabolism* 102 (8): 3040–49.
637 McQuillan, Ruth, Harry Campbell, Veronique Vitart, and James Wilson. 2009. *Homozygosity,*

*Inbreeding and Health in European Populations*.
McQuillan, Ruth, Niina Eklund, Nicola Pirastu, Maris Kuningas, Brian P. McEvoy, Tõnu Esko, Tanguy Corre, et al. 2012. "Evidence of Inbreeding Depression on Human Height." *PLoS Genetics* 8 (7). https://doi.org/10.1371/journal.pgen.1002655.
Melhem, Nadine M., Cong Lu, Cara Dresbold, Frank A. Middleton, Lambertus Klei, Shawn Wood, Stephen V. Faraone, et al. 2014. "Characterizing Runs of Homozygosity and Their Impact on Risk for Psychosis in a Population Isolate." *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics* 165B (6): 521.
Mooney, Jazlyn A., Abigail Yohannes, and Kirk E. Lohmueller. 2021. "The Impact of Identity by Descent on Fitness and Disease in Dogs." *Proceedings of the National Academy of Sciences of the United States of America* 118 (16). https://doi.org/10.1073/pnas.2019116118.
Moreno-Grau, Sonia, Maria Victoria Fernández, Itziar de Rojas, Pablo Garcia-González, Isabel Hernández, Fabiana Farias, John P. Budde, et al. 2021. "Long Runs of Homozygosity Are Associated with Alzheimer's Disease." *Translational Psychiatry* 11 (1): 1–12.
Parker, H. G., B. M. VonHoldt, P. Quignon, E. H. Margulies, S. Shao, D. S. Mosher, T. C. Spady, et al. 2009. "An Expressed fgf4 Retrogene Is Associated with Breed-Defining Chondrodysplasia in Domestic Dogs." 10.1126/science.1173275.
Pemberton, Trevor J., Devin Absher, Marcus W. Feldman, Richard M. Myers, Noah A. Rosenberg, and Jun Z. Li. 2012. "Genomic Patterns of Homozygosity in Worldwide Human Populations." *American Journal of Human Genetics* 91 (2): 275–92.
Peritogiannis, Vaios, Angeliki Ninou, and Maria Samakouri. 2022. "Mortality in Schizophrenia-Spectrum Disorders: Recent Advances in Understanding and Management." *HealthcarePapers* 10 (12). https://doi.org/10.3390/healthcare10122366.
Plassais, Jocelyn, Jaemin Kim, Brian W. Davis, Danielle M. Karyadi, Andrew N. Hogan, Alex C. Harris, Brennan Decker, Heidi G. Parker, and Elaine A. Ostrander. 2019. "Whole Genome Sequencing of Canids Reveals Genomic Regions under Selection and Variants Influencing Morphology." *Nature Communications* 10 (1): 1–14.
Plummer, Martyn, Nicky Best, Kate Cowles, and Karen Vines. 2006. "CODA: Convergence Diagnosis and Output Analysis for MCMC." *R News* 6 (1): 7–11.
Purfield, Deirdre C., Sinead McParland, Eamon Wall, and Donagh P. Berry. 2017. "The Distribution of Runs of Homozygosity and Selection Signatures in Six Commercial Meat Sheep Breeds." *PloS One* 12 (5). https://doi.org/10.1371/journal.pone.0176780.
Quinlan, Aaron R. 2014. "BEDTools: The Swiss-Army Tool for Genome Feature Analysis." *Current Protocols in Bioinformatics / Editoral Board, Andreas D. Baxevanis ... [et Al.]* 47 (September): 11.12.1–34.
Sams, Aaron J., and Adam R. Boyko. 2019. "Fine-Scale Resolution of Runs of Homozygosity Reveal Patterns of Inbreeding and Substantial Overlap with Recessive Disease Genotypes in Domestic Dogs."
Selionova, Marina, Magomet Aibazov, Tatiana Mamontova, Victor Malorodov, Alexander Sermyagin, Natalia Zinovieva, and Ahmed A. Easa. 2022. "Genome-Wide Association Study of Live Body Weight and Body Conformation Traits in Young Karachai Goats." *Small Ruminant Research* 216 (November). https://doi.org/10.1016/j.smallrumres.2022.106836.
Shapses, Sue A., and Deeptha Sukumar. 2012. "Bone Metabolism in Obesity and Weight Loss." *Annual Review of Nutrition* 32 (August): 287–309.
Sofer, Tamar, Xiuwen Zheng, Cecelia A. Laurie, Stephanie M. Gogarten, Jennifer A. Brody, Matthew P. Conomos, Joshua C. Bis, et al. 2021. "Variant-Specific Inflation Factors for Assessing Population Stratification at the Phenotypic Variance Level." *Nature Communications* 12 (1): 1–14.
Sollis, Elliot, Abayomi Mosaku, Ala Abid, Annalisa Buniello, Maria Cerezo, Laurent Gil, Tudor

Groza, et al. 2023. “The NHGRI-EBI GWAS Catalog: Knowledgebase and Deposition Resource.” *Nucleic Acids Research* 51 (D1): D977–85.

Sutter, Nathan B., Carlos D. Bustamante, Kevin Chase, Melissa M. Gray, Keyan Zhao, Lan Zhu, Badri Padhukasahasram, et al. 2007. “A Single IGF1 Allele Is a Major Determinant of Small Size in Dogs.”

Swinford, Natalie, Brenna Gallagher, Jacob Sheehama, Meng Lin, Sean Prall, Brooke Scelza, and Brenna M. Henn. 2023. “Examination of Runs of Homozygosity in Relation to Height in an Endogamous Namibian Population.” *American Journal of Biological Anthropology* 180 (1): 207.

Szpiech, Zachary A., Alexandra Blant, and Trevor J. Pemberton. 2017. “GARLIC: Genomic Autozygosity Regions Likelihood-Based Inference and Classification.”

Tachmazidou, Ioanna, Dániel Süveges, Josine L. Min, Graham R. S. Ritchie, Julia Steinberg, Klaudia Walter, Valentina Iotchkova, et al. 2017. “Whole-Genome Sequencing Coupled to Imputation Discovers Genetic Signals for Anthropometric Traits.” *American Journal of Human Genetics* 100 (6): 865–84.

Thompson, Elizabeth A. 2013. “Identity by Descent: Variation in Meiosis, Across Genomes, and in Populations.” *Genetics* 194 (2): 301.

Trujillo-Viera, Jonathan, Rabih El-Merahbi, Vanessa Schmidt, Till Karwen, Angel Loza-Valdes, Akim Strohmeyer, Saskia Reuter, et al. 2021. “Protein Kinase D2 Drives Chylomicron-Mediated Lipid Transport in the Intestine and Promotes Obesity.” *EMBO Molecular Medicine* 13 (5): e13548.

Turner, Stephen D. 2018. “Qqman: An R Package for Visualizing GWAS Results Using Q-Q and Manhattan Plots.” *Journal of Open Source Software* 3 (25): 731.

Urfer, Silvan R., Mansen Wang, Mingyin Yang, Elizabeth M. Lund, and Sandra L. Lefebvre. 2019. “Risk Factors Associated with Lifespan in Pet Dogs Evaluated in Primary Care Veterinary Hospitals.” *Journal of the American Animal Hospital Association* 55 (3): 130–37.

Vaysse, Amaury, Abhirami Ratnakumar, Thomas Derrien, Erik Axelsson, Gerli Rosengren Pielberg, Snaevar Sigurdsson, Tove Fall, et al. 2011. “Identification of Genomic Regions Associated with Phenotypic Variation between Dog Breeds Using Selection Mapping.” *PLoS Genetics* 7 (10): e1002316.

Vilà, Carles, Peter Savolainen, Jesús E. Maldonado, Isabel R. Amorim, John E. Rice, Rodney L. Honeycutt, Keith A. Crandall, Joakim Lundeberg, and Robert K. Wayne. 1997. “Multiple and Ancient Origins of the Domestic Dog.” *Science*, June. https://doi.org/10.1126/science.276.5319.1687.

Virakul, Sita, Poorichaya Somparn, Trairak Pisitkun, Peter J. van der Spek, Virgil A. S. H. Dalm, Dion Paridaens, P. Martin van Hagen, Nattiya Hirankarn, Tanapat Palaga, and Willem A. Dik. 2020. “Integrative Analysis of Proteomics and DNA Methylation in Orbital Fibroblasts From Graves’ Ophthalmopathy.” *Frontiers in Endocrinology* 11: 619989.

vonHoldt, Bridgett M., John P. Pollinger, Kirk E. Lohmueller, Eunjung Han, Heidi G. Parker, Pascale Quignon, Jeremiah D. Degenhardt, et al. 2010. “Genome-Wide SNP and Haplotype Analyses Reveal a Rich History Underlying Dog Domestication.” *Nature* 464 (7290): 898–902.

Warrington, Nicole M., Robin N. Beaumont, Momoko Horikoshi, Felix R. Day, Øyvind Helgeland, Charles Laurin, Jonas Bacelis, et al. 2019. “Maternal and Fetal Genetic Effects on Birth Weight and Their Relevance to Cardio-Metabolic Risk Factors.” *Nature Genetics* 51 (5): 804–14.

Wu, Wentao, Daning Li, Xiaojie Feng, Fanfan Zhao, Chengzhuo Li, Shuai Zheng, and Jun Lyu. 2021. “A Pan-Cancer Study of Selenoprotein Genes as Promising Targets for Cancer Therapy.” *BMC Medical Genomics* 14 (1): 78.

Yang, Qianyong, Hao Chen, Junhua Ye, Chenlong Liu, Rongxing Wei, Congying Chen, and Lusheng Huang. 2019. “Genetic Diversity and Signatures of Selection in 15 Chinese

Indigenous Dog Breeds Revealed by Genome-Wide SNPs." *Frontiers in Genetics* 10 (November): 489469.

Yang, Wanneng, Zilong Guo, Chenglong Huang, Lingfeng Duan, Guoxing Chen, Ni Jiang, Wei Fang, et al. 2014. "Combining High-Throughput Phenotyping and Genome-Wide Association Studies to Reveal Natural Genetic Variation in Rice." *Nature Communications* 5 (1): 1–9.

Yengo, Loïc, Sailaja Vedantam, Eirini Marouli, Julia Sidorenko, Eric Bartell, Saori Sakaue, Marielisa Graff, et al. 2022. "A Saturated Map of Common Genetic Variants Associated with Human Height." *Nature* 610 (7933): 704–12.

Yordy, Jennifer, Cornelia Kraus, Jessica J. Hayward, Michelle E. White, Laura M. Shannon, Kate E. Creevy, Daniel E. L. Promislow, and Adam R. Boyko. 2019. "Body Size, Inbreeding, and Lifespan in Domestic Dogs."

Yuan, Jie, Xueqiang Wang, Yan Zhao, Najeeb Ullah Khan, Zhiqiang Zhao, Yanhong Zhang, Xiaorong Wen, Fusen Tang, Fengbin Wang, and Zichao Li. 2020. "Genetic Basis and Identification of Candidate Genes for Salt Tolerance in Rice by GWAS." *Scientific Reports* 10 (1): 9958.

Zhao, Weizhu, Haiying Yang, Jie Chai, and Ligang Xing. 2021. "RUNX2 as a Promising Therapeutic Target for Malignant Tumors." *Cancer Management and Research* 13 (March): 2539–48.